Review of Bharathan et al.: Enhancing GOSAT methane observations over the high latitudes using a multi-objective genetic algorithm optimisation approach, submitted to Atmospheric Measurement Techniques

5 May 2026

---

**Review Summary**

The study by Bharathan et al. utilises ground-based observations of total column methane from the high-latitude TCCON to optimise the post-processing filtering for satellite-based methane retrievals from GOSAT. The optimization is done using a multi-objective genetic algorithm. As a result, they are able to significantly increase the number of high-latitude observations by a reasonable compromise in the RMSE. The high-latitude-optimised product is also shown to have a good global performance.

In our review, we find this study and the high-latitude-optimised dataset highly useful for applications such as inverse modelling or satellite-data-driven approaches. An increase in the data amounts during the data-poor seasons can be highly significant when optimizing for fluxes in satellite-based inversions, also beyond the Arctic (though not shown in the study). This data product is therefore of high interest for the community. This paper describing the generation and validation of the GOSAT data product is appropriate for AMT, and we are happy to recommend it for publication. However, we have two broader questions, multiple minor comments and questions, and some technical corrections that would need to be addressed before publishing should be considered. We have detailed these below. In addition, we recommend making the postprocessed end product publicly available when publishing the paper.

**Major comments**

1) We recommend that the authors would more clearly describe the convergence of the optimisation and the end of generations. Currently, there is no description of the stopping criterion for the algorithm used in the study, though it appears to be a predetermined number of generations, i.e., 100. It would be good to add in the paper more discussion on how to ensure convergence has been reached, whether and why subsequent generations result in improvements, and how the limiting 100 generations has been selected. Since the solutions seem to stabilize after 15-20 generations, and the population becomes filled with non-dominated individuals, is there a danger that the solutions are crowded, or is this fast convergence desired behaviour?

2) a) The optimisation is based on the RMSE between high-latitude TCCON and the GOSAT observations. We recommend adding in a short description on the calculation of this RMSE, including the co-location of GOSAT and TCCON observations. b) Evaluating the optimized filtering using TCCON data (Sect. 4.4) is somewhat circular, considering the role of TCCON data in the optimization

process. We recommend adding in a short discussion on how the authors justify their approach, instead of utilizing other available options (such as comparisons against the EM27/SUN instruments).

**Minor comments**

1) Lines 20-21: Arctic warming has been shown to be three to four times faster than the global average (Rantanen et al., 2022).
2) Line 24: what is meant by the exposure of organic matter? This may be largely focusing on the thawing of permafrost but Arctic methane emissions are also dependent on seasonal frost and snow.
3) Lines 34-36: challenging weather conditions are only part of the explanation for the lack of Arctic measurements. Large regions of the Arctic are not accessible due to geopolitics and/or lacking infrastructure, such as roads or permanently manned stations.
4) Line 45: Low solar angle → Either write "low solar elevation angle" or "large solar zenith angle" to avoid confusion.
5) Line 49: "Ice" → "snow and ice" because the same applies to snow. Although snow is in practice ice, this makes the message more clear.
6) Line 71 (also elsewhere): There are multiple terminologies in the paper to describe $XCH_4$. Please choose one and use it systematically. Also, 4 should be in subscript, and the X should not be italicized (e.g. line 179).
7) Paragraph starting on line 73: this is interesting and relevant background, and directly ties to the motivation of this paper. We recommend to extend this part with a description of the current work and how it is related to the previous work (perhaps move the last sentence from the previous paragraph here).
8) Figure 1: What is the motivation for showing this figure? For a high-latitude-focused paper we would recommend considering a polar projection, or at least drawing a 50 N line on the global map for pointing out the focus region. In addition, the caption reads that gridded data are shown; however, the data seem to be Level 2 retrievals (please check). The text says "mean GOSAT UoL Proxy XCH4" – please check and harmonise with what is shown in Fig. 1 (mean gridded values or L2 observations?).
9) Line 113. GOSAT XCH4 → please specify the product name
10) Line 115. "Low solar zenith angles" should be low solar elevation angles or large solar zenith angles, please correct.
11) Line 116. What are the extreme atmospheric conditions, please specify.
12) Line 117. It is interesting that about a third of cloud-cleared observations are lost during the quality filtering process. Is this a global mean value? Is it about the same for high latitudes – we'd assume larger? Is there seasonal variability? This could be described with a couple additional sentences since it partly motivates the actual work.
13) Figure 2. We suggest removing Figure 2b which does not provide additional information considering Fig. 2a. A few latitude lines might be a useful addition in Fig. 2a. Figure 2b could potentially be replaced by a polar view of the data count; we leave this up to the authors' consideration.

14) Line 171. Are these fit individuals passed on to the next generation as-is, or are these used as parents to generate the next generation?
15) Section 3.1: How were the eight filter values selected? Some are obvious (e.g. high chi-square) but why a lower limit of chi-square is needed, and why low thresholds of $XCH_4$ and $XCO_2$ but not high? Are the "uncertainty values" the statistical uncertainties from the retrieval process? Did the authors experiment with any other filter values?
16) Table 1: We suggest to remove this table entirely as it does not bring any added information with respect to the text.
17) Figure 3: Since the focus of this study is on the high latitudes, we would suggest only showing the region of interest in this figure. Also, both the caption and text describe the figure as showing averages of the filter parameters, but there is no mention of how the data have been averaged (spatio-temporally?); the data also seem to correspond to Level 2 retrievals rather than averages, as in Fig. 1, please check.
18) Line 192-195 (and elsewhere). Here, within a few lines, the word "solution" appears in three different contexts as equal to individual, chromosome, and offspring. We would suggest sticking to a limited set of words throughout most of the manuscript, instead of giving multiple options each time.
19) Line 210. The sentence here starts with "similarly", but it's unclear what the mutation process described is similar to. Perhaps "additionally" would work better, if mutation takes place after crossover. Are the mutation performed on solutions created through crossover, or are these completely separate processes, as might be interpreted from Table 2. Please clarify.
20) Line 215. What is the difference between ranks and (pareto) fronts? Please clarify. If they are essentially the same thing, consider using only one of these terms.
21) Figure 5, caption. C: 15 generations; should this be 20? Please check.
22) Figure 6: Since all the panels have the same colour scale, please use a single colour bar for the figure.
23) Figure 7c: are there any negative values? It is not obvious from the figure. If there are, should the color scale be reconsidered or saturated to better show these values?
24) Figure 7: There seem to be quite a lot more observations also over the tropical / subtropical oceans. We recommend to point this out in the text and discuss why this is the case.
25) Figure 8b: the y scale reads fractional increase and the caption says total; please harmonise.
26) Line 267: emission flux sounds redundant; we suggest to use emissions
27) Line 268: not only permafrost; we suggest to also include wetlands
28) Section 4.4: it would be helpful to the readers if the authors added one representative example of a high-latitude time series with the TCCON data, "default" GOSAT Proxy $XCH_4$ data, and added data with the new postprocessing.
29) Line 308: The citation given here is to a GHG-CCI data set, not the ESA GHG-CCI User Requirements Document. Please check and correct.
30) Lines 295-296 and references. Please correct the citations featuring non-standard English characters. Please check and correct the entire reference list

for similar or any other issues (e.g. no authors on the first listed reference, the editors of the IPCC 2023 report (line 417)).

31) Data availability: We highly recommend to make the postprocessed data version publicly available (or adding a new filtering flag to the existing data product) to facilitate its use.
32) Appendix A: Few sentences of text could be added in the Appendix.

**Technical corrections**

Line 1. ”Gases” missing from GOSAT
Line 7. Unclear sentence; perhaps remove ”disproportionately affects”
Line 18. Expand abbreviations (CO2M etc.)
Line 36. its → it is
Line 40 (and also elsewhere). Parentheses should be written like this when using LaTeX: \citep[GOSAT][]{Kuze et al. 2009} to avoid subsequent parentheses.
Line 42. Long-periods → can be removed
Line 42. Add space after records
Line 70. Remove comma after of
Line 83. Reference (Kuz, 2014) is missing
Line 107. Add space after observations
Line 115. Nights → night
Line 118. Missing the before complete omission
Line 125-127 (and elsewhere). The N should not be italicized when denoting latitude
Line 133. time-periods → time periods
Line 133. carry should be carries
Line 174. crowded → “getting crowded” or “crowding”
Line 175 (and elsewhere). The abbreviation “NSGA II” was used on line 70, but here it is written as “NSGA-II”. Please use one form consistently.
Line 187. You start two consecutive sentences with so here. Maybe remove the first to avoid repetition.
Line 209. 200-members; please remove the dash
Line 213. Remove comma after mentioned
Line 224. carry → carries
Line 226. forms → form
Line 240. 6. should be 6
Line 248. Missing of after number
Line 251. Correct reference to Table 3, not Table 2.
Figure 7 caption. standared → standard. Also remove comma before c) and correct Delta to the Greek letter.
Line 266. Missing a before severe
Line 287. Capitalise Appendix
Figure 9 caption. High → high
Line 309. Remove such as, since you go on to list all the metrics in Table 4 (or rephrase the sentence)
Line 323. Quality-filters → quality filters

## References

Rantanen, M., Karpechko, A.Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T. and Laaksonen, A., 2022. The Arctic has warmed nearly four times faster than the globe since 1979. *Communications earth & environment*, *3*(1), p.168.